Nonlinear Hilbert Adjoints

Gray, W. Steven; Scherpen, Jacquelien M.A.

Published in:
Proceedings of the 2001 American Control Conference

Publication date:
2001

Citation for published version (APA):
Gray, W. S., & Scherpen, J. M. A. (2001). Nonlinear Hilbert Adjoints: Properties and Applications to Hankel Singular Value Analysis. In Proceedings of the 2001 American Control Conference (Vol. 5, pp. 3582-3587). University of Groningen, Research Institute of Technology and Management.

Proceedings of the American Control Conference
Arlington, VA June 25-27, 2001

Nonlinear Hilbert Adjoints: Properties and Applications to Hankel Singular Value Analysis

W. Steven Gray

Department of Electrical and Computer Engineering
Old Dominion University
Norfolk, Virginia 23529-0246
U.S.A.
gray@ece.odu.edu

Jacquelien M. A. Scherpen

Delft University of Technology
Fac. of Inf. Techn. & Syst.
Department of Electrical Engineering
P.O. Box 5031, 2600 GA Delft
The Netherlands
J.M.A.Scherpen@et.tudelft.nl

Abstract

The notion of an adjoint operator for a nonlinear mapping has few interpretations in the literature. In this paper a new nonlinear Hilbert adjoint operator is proposed. It is shown to unite several existing concepts and provides an essential tool for singular value analysis of nonlinear Hankel operators.

1. Introduction

Once one departs from the context of linear operators, there are very few extensions of the adjoint operator definition. Furthermore, it can not be assumed a priori that the existing notions are in any way directly related. For example, in [2] the notion of an adjoint map is defined in terms of a dual map on a topological vector space. This idea is distinct from the adjoint map that appears in [4, 14] which employs the Gâteaux derivative of the operator when it is well defined. In a nonlinear state space context, the adjoint system has appeared in [5], but only recently has it been given an input-output interpretation using a nonlinear Hilbert adjoint operator [7, 8]. This latter concept first appeared in an abstract setting in [10, 16] mainly to address the open problem of understanding how to relate the state space notion of singular value functions due to Scherpen [15] to the nonlinear Hankel operator extension. In this paper the basic objective is to fully develop the idea of a nonlinear Hilbert adjoint and to further illustrate its usefulness in Hankel singular value analysis.

2. Nonlinear Hilbert Adjoint Operators

In the most general setting, let F be a topological vector space over $\mathbb{R}$ with dual space F'. Let E be a nonempty set, and $\mathcal{A}$ a collection of nonempty subsets of E. Let E^β be a linear space of real-valued functions x^β on E with the property that the restriction x^β_A to every $A \in \mathcal{A}$ is bounded. A mapping $\mathcal{T} : E \rightarrow F$ is called $\mathcal{A}$*-bounded* if $\mathcal{T}$ maps the sets of $\mathcal{A}$ into bounded subsets of F. For any $\mathcal{A}$ -bounded mapping $\mathcal{T} : E \rightarrow F$, the *dual map* of $\mathcal{T}$ is defined as

$$\begin{aligned} \mathcal{T}' &: F' \rightarrow E^\beta \\ &: y' \rightarrow (\mathcal{T}'(y'))(u) = (y' \circ \mathcal{T})(u), \quad \forall u \in E \end{aligned}$$

(see, for example, [2]). Now if F is a Hilbert space with an inner product $\langle \cdot, \cdot \rangle_F$ then it follows from the Riesz Lemma that for any $y' \in F'$ there exists a unique $y \in F$ such that $y'(\cdot) = \langle y, \cdot \rangle_F$. Hence one can write the identity

$$(\mathcal{T}'(y'))(u) = \langle y, \mathcal{T}(u) \rangle_F, \quad \forall u \in E.$$

If, in addition, E is an inner product space with inner product $\langle \cdot, \cdot \rangle_E$ and $y \in F$ is fixed, then the problem is to determine a corresponding $\tilde{u}_y \in E$ such that

$$\langle \mathcal{T}(u), y \rangle_F = \langle u, \tilde{u}_y \rangle_E, \quad \forall u \in E. \tag{1}$$

If $\mathcal{T}$ were a linear operator then such an $\tilde{u}_y$ is known to always exist and be unique, i.e., $\tilde{u}_y = \mathcal{T}^*(y)$, where $\mathcal{T}^*$ is the Hilbert adjoint of $\mathcal{T}$. But in this more general context, the existence and uniqueness of $\tilde{u}_y$ are not automatic. In fact, the identity (1) is meaningful in most cases only when $\tilde{u}_y$ is also a function of u. (Defining the domain of $\mathcal{T}^*$ to have the form $F \times E$ also agrees with the state space notion of adjoint *systems* based on the Hamiltonian extension given in [5, 17].) So in this context, consider the following definition.

Definition 2.1 *Given two Hilbert spaces E and F, an operator $\mathcal{T} : E \mapsto F$ has a global* ***nonlinear Hilbert adjoint*** *when there exists an operator $\mathcal{T}^* : F \times E \rightarrow E$ such that*

$$\langle \mathcal{T}(u), y \rangle_F = \langle u, \mathcal{T}^*(y, u) \rangle_E, \quad \forall u \in E, \quad \forall y \in F, \tag{2}$$

where $\mathcal{T}^*(y, u)$ is linear in y.

The above definition is more general than the definition of an adjoint operator given in [4], where the identity (2) is only required to hold when $y = u$. Our interest in studying adjoint operators originated in the study of singular value structures, which implies in the above definition that $y = \mathcal{T}(u)$ should be admissible. The adjoint definition of [4] is too limited for this purpose. Also, there appears to be some parallel development of Hilbert adjoints in the Russian literature which recently became more widely available in [14]. It should be noted that Definition 2.1 is slightly different from that which appeared in [10, 16] since here linearity in y is an additional requirement. It seems rather natural in light of the bilinearity of inner products. But linearity in y is not automatic from (2) because it is often the case that there exists a collection of nontrivial mappings (linear and nonlinear in y) of the form $\mathcal{B} : F \times E \mapsto E$ such that

$\langle u, \mathcal{B}(y,u)\rangle_E = 0, \forall u \in E, \forall y \in F$. In which case, any adjoint mapping $\mathcal{T}^*$ is not uniquely defined since $\mathcal{T}^* + \mathcal{B}$ will also satisfy equation (2). In these circumstances, an adjoint operator should be viewed as a member of an equivalence class where two such operators $\mathcal{T}^*$ and $\mathcal{T}^{*'}$ are equivalent if

$$\langle u, \mathcal{T}^*(y,u)\rangle_E = \langle u, \mathcal{T}^{*'}(y,u)\rangle_E, \quad \forall u \in E, \quad \forall y \in F. \tag{3}$$

A shorthand notation for (3) is simply $\mathcal{T}^*(y,u) = \mathcal{T}^{*'}(y,u)$. Thus, any equality involving adjoint operators really means that both expressions belong to the same equivalence class. (See [11] for analysis and examples closely related to this issue.)

It is not necessary in many applications to have a globally defined $\mathcal{T}^*$. The following theorem gives a sufficient condition for the existence of a locally defined adjoint operator.

Theorem 2.1 *Suppose H_1 and H_2 are two Hilbert spaces and $U \subset H_1$ is any convex neighborhood of 0. Let $\mathcal{T} : U \mapsto H_2$ be a continuously Fréchet differentiable mapping on U such that $\mathcal{T}(0) = 0$. Then the mapping*

$$\mathcal{T}^*(y,u) = \int_0^1 (D\mathcal{T}(tu))^*(y)\, dt,$$

where $D\mathcal{T}$ denotes the Fréchet derivative of $\mathcal{T}$, is a suitable Hilbert adjoint of $\mathcal{T}$ on $H_2 \times U$.

Proof: For any $y \in H_2$, define the scalar-valued mapping on U:

$$L_y(u) = \langle \mathcal{T}(u), y\rangle_{H_2} \equiv \langle u, \mathcal{T}^*(y,u)\rangle_{H_1}.$$

Next observe that for any fixed $u \in U$ and $t \in [0,1]$ it follows that

$$\begin{aligned} DL_y(tu)(\xi) &= \langle D\mathcal{T}(tu)(\xi), y\rangle_{H_2} \\ &= \langle \xi, (D\mathcal{T}(tu))^*(y)\rangle_{H_1}, \quad \forall \xi \in H_1. \end{aligned}$$

Thus,

$$\begin{aligned} L_y(u) &= \int_0^1 (DL_y(tu))(u)\, dt \\ &= \int_0^1 \langle u, (D\mathcal{T}(tu))^*(y)\rangle_{H_1}\, dt \\ &= \langle u, \int_0^1 (D\mathcal{T}(tu))^*(y)\, dt\rangle_{H_1}, \end{aligned}$$

and the conclusion follows directly. ∎

Observe that in this form above, $\mathcal{T}^*(y,u)$ is linear in y since $(D\mathcal{T}(tu))^*(\cdot)$ is the adjoint of a linear operator, i.e., the familiar Hilbert adjoint. Thus, it is also immediate that $\mathcal{T}^*(0,u) = 0$, $\forall u \in U$.

Example 2.1 For any finite $T > 0$ and positive integer m, the Banach space $L_4^m[0,T]$ can be viewed as a convex open subset of $L_2^m[0,T]$ containing the zero function. With $U = L_4^m[0,T]$, the mapping

$$\begin{aligned} \mathcal{T} &: U \mapsto L_2[0,T] \\ &: u \mapsto u^T u \end{aligned}$$

is then well defined, continuously Fréchet differentiable, and satisfies the identity $\mathcal{T}(0) = 0$. One form of the adjoint operator can be immediately extracted using the definition. Specifically,

$$\begin{aligned} \langle \mathcal{T}(u), y\rangle_{L_2} &= \int_0^T u^T(\tau)u(\tau)y(\tau)\, d\tau, \\ &= \langle u, uy\rangle_{L_2^m}, \end{aligned}$$

and thus, $\mathcal{T}^*(y,u) = uy$. This same adjoint form can also be computed using Theorem 2.1:

$$\begin{aligned} D\mathcal{T}(u) &= 2u^T \\ (D\mathcal{T}(tu))^*(y) &= 2tuy \\ \mathcal{T}^*(y,u) &= \int_0^1 (D\mathcal{T}(tu))^*(y)\, dt = uy. \end{aligned}$$

□

Example 2.2 Consider an Hammerstein integral operator defined on a set $U \subset L_2^m[0,\infty)$:

$$\begin{aligned} \mathcal{S} &: U \mapsto L_2^p[0,\infty) \\ &: u \mapsto \int_0^\infty K(\tau,s)f(u(s))\, ds, \end{aligned}$$

where K is a suitable continuous kernel function, and each component function of f is C^1 with $f(0) = 0$. Then applying Theorem 2.1 it follows that

$$\begin{aligned} (D\mathcal{S}(u))(\xi) &= \int_0^\infty K(\tau,s)\frac{df}{du}(u(s))\xi(s)\, ds \\ (D\mathcal{S}(u))^*(y) &= \int_0^\infty \frac{df}{du}^T(u(s))K^T(\tau,s)y(\tau)\, d\tau \\ \mathcal{S}^*(y,u) &= \int_0^\infty \underbrace{\left[\int_0^1 \frac{df}{du}(tu(s))\, dt\right]^T}_{F(u(s))} K^T(\tau,s)y(\tau)\, d\tau \\ &= F^T(u)\mathcal{S}_{\mathrm{L}}^*(y), \end{aligned}$$

where the matrix-valued function $F(\cdot)$ satisfies the identity $f(x) = F(x)x$ on a convex neighborhood of 0, and $\mathcal{S}_{\mathrm{L}}^*$ denotes the usual adjoint operator for the linear integral operator with kernel K. For example, the SISO FM system

$$\mathcal{S}_{\mathrm{FM}}(u) = \frac{1}{\pi}\int_0^\infty e^{A(\tau - s)} sin(\pi\, u(s))\, ds$$

gives

$$\begin{aligned} \mathcal{S}_{\mathrm{FM}}^*(y,u) &= sinc(u)\int_0^\infty e^{A^T(\tau-s)}y(\tau)\, d\tau \\ &= sinc(u)\mathcal{S}_{\mathrm{L}}^*(y). \end{aligned}$$

□

Consider any normed set of linear operators B defined on $L_2[0,\infty)$ as a Banach algebra with composition product $(\mathcal{S},\mathcal{T}) \mapsto \mathcal{S}\mathcal{T}$. B is said to constitute a C^*-algebra if it is equipped with an *adjoint map* (or *involution*) $\mathcal{T} \mapsto \mathcal{T}^*$ such that for all $\mathcal{S},\mathcal{T} \in B$ and any $\alpha \in \mathbb{R}$, the following properties are satisfied:

i. (linearity) $(\alpha\mathcal{S} + \mathcal{T})^* = \alpha\mathcal{S}^* + \mathcal{T}^*$;

ii. (product-reversal) $(\mathcal{S}\mathcal{T})^* = \mathcal{T}^*\mathcal{S}^*$;

iii. (double adjoint) $(\mathcal{T}^*)^* = \mathcal{T}$; and

iv. (C^*-identity) $\|\mathcal{T}\|^2 = \|\mathcal{T}^*\mathcal{T}\|$.

We next provide the appropriate extensions of these fundamental properties for the nonlinear Hilbert adjoint. The linearity property (i) is an immediate result which follows from the bilinearity of the inner product and the interpretation that equality here implies belonging to the same equivalence class. In order to address the product-reversal property (ii), one must first define the sense in which operators can be composed when adjoint operators are present. The situation is more complicated than the familiar case since the domain of an adjoint operator is not simply the codomain of the original operator. For example, consider the Hilbert spaces H_i, $i = 1, 2, 3$, the operators

$$\begin{array}{ll} \mathcal{T}: H_1 \mapsto H_2 & \mathcal{S}: H_2 \mapsto H_3 \\ \quad : u \mapsto w & \quad : w \mapsto y \end{array}$$

and the corresponding adjoints

$$\begin{array}{ll} \mathcal{T}^*: H_2 \times H_1 \mapsto H_1 & \mathcal{S}^*: H_3 \times H_2 \mapsto H_2 \\ \quad : (w, u) \mapsto \bar{u} & \quad : (y, w) \mapsto \bar{w}. \end{array}$$

Clearly the composition and its adjoint

$$\begin{array}{ll} \mathcal{ST}: H_1 \mapsto H_3 & (\mathcal{ST})^*: H_3 \times H_1 \mapsto H_1 \\ \quad : u \mapsto y & \quad : (y, u) \mapsto \bar{u}. \end{array}$$

are well defined, but no *direct* composition like $\mathcal{T}^*\mathcal{T}$ or $\mathcal{T}^*\mathcal{S}^*$ is possible as in the classic setting. Still some *formal* compositions can be defined which have great utility in a variety of situations.

Definition 2.2 *Let H_i, $i = 1, 2, 3$, be a collection of Hilbert spaces. Assume $\mathcal{T} : H_1 \mapsto H_2$ and $\mathcal{S} : H_2 \mapsto H_3$ are two operators with well-defined adjoint operators. Define the following operator products:*

$$\begin{array}{lll} (\mathcal{S}^*\mathcal{T})_1 & : & H_1 \times H_2 \mapsto H_2 \quad [\text{when } H_2 = H_3] \\ & : & (u, w) \mapsto \mathcal{S}^*(\mathcal{T}(u), w) \\ & & \\ (\mathcal{S}^*\mathcal{T})_2 & : & H_3 \times H_1 \mapsto H_1 \\ & : & (y, u) \mapsto \mathcal{S}^*(y, \mathcal{T}(u)). \end{array}$$

Theorem 2.2 *(product-reversal) Let H_i, $i = 1, 2, 3$, be a collection of Hilbert spaces. Assume $\mathcal{T} : H_1 \mapsto H_2$ and $\mathcal{S} : H_2 \mapsto H_3$ are two operators with well-defined adjoint operators. Then the following identity holds:*

$$(\mathcal{ST})^* = (\mathcal{T}^*(\mathcal{S}^*\mathcal{T})_2)_1.$$

Proof: The claim follows straightforwardly from the defining property (2). Observe that for any $(y, u) \in H_3 \times H_1$:

$$\begin{aligned} \langle u, (\mathcal{ST})^*(y, u)\rangle_{H_1} &= \langle \mathcal{ST}(u), y\rangle_{H_3} \\ &= \langle \mathcal{T}(u), \mathcal{S}^*(y, \mathcal{T}(u))\rangle_{H_2} \\ &= \langle u, \mathcal{T}^*(\mathcal{S}^*(y, \mathcal{T}(u)), u)\rangle_{H_1} \\ &= \langle u, (\mathcal{T}^*(\mathcal{S}^*\mathcal{T})_2)_1(y, u)\rangle_{H_1}. \end{aligned}$$

■

In order to compute adjoints of general adjoint operators for the double adjoint property (iii), the concept of a partial adjoint operator is needed. The idea is based on a direct generalization of identity (2).

Definition 2.3 *For any set of Hilbert spaces H_i, $i = 1, \ldots, m+1$ and an operator*

$$\begin{array}{lll} \mathcal{U} & : & H_1 \times \ldots \times H_m \mapsto H_{m+1} \\ & : & (u_1 \ldots, u_m) \mapsto y, \end{array} \tag{4}$$

*a **jth partial adjoint of $\mathcal{U}$** is any mapping of the form*

$$\mathcal{U}^{*_j} : H_{m+1} \times H_1 \ldots \times H_m \mapsto H_j,$$

where

$$\langle \mathcal{U}(u_1, \ldots, u_m), y\rangle_{H_{m+1}} = \langle u_j, \mathcal{U}^{*_j}(y, u_1, \ldots, u_m)\rangle_{H_j}$$

for all $u_i \in H_i$, $i = 1, \ldots, m$, and $y \in H_{m+1}$. These definitions produce the following double adjoint identities.

Theorem 2.3 *(double adjoints) Let H_1 and H_2 be two Hilbert spaces and $\mathcal{T} : H_1 \mapsto H_2$ be an operator with a well defined adjoint. Then it follows that*

$$\begin{aligned} (\mathcal{T}^*)^{*_1}(\bar{u}, y, u)|_{\bar{u}=u} &= \mathcal{T}(u) \\ (\mathcal{T}^*)^{*_2}(\bar{u}, y, u)|_{\bar{u}=u} &= \mathcal{T}^*(y, u) \end{aligned}$$

for all $u \in H_1, y \in H_2$, assuming all the partial adjoints exist.

Proof: With respect to the first identity, observe that the first partial adjoint of $\mathcal{T}^*(y, u)$ fulfills

$$\begin{aligned} \langle y, (\mathcal{T}^*)^{*_1}(\bar{u}, y, u)\rangle|_{\bar{u}=u} &= \langle \mathcal{T}^*(y, u), \bar{u}\rangle|_{\bar{u}=u} \\ &= \langle y, \mathcal{T}(u)\rangle. \end{aligned}$$

For the second partial adjoint of $\mathcal{T}^*(y, u)$,

$$\begin{aligned} \langle u, (\mathcal{T}^*)^{*_2}(\bar{u}, y, u)\rangle|_{\bar{u}=u} &= \langle \mathcal{T}^*(y, u), \bar{u}\rangle|_{\bar{u}=u} \\ &= \langle u, \mathcal{T}^*(y, u)\rangle. \end{aligned}$$

■

One application of this theorem is in regards to testing for self-adjointness.

Definition 2.4 *Let H be a Hilbert space and $\mathcal{S} : H \mapsto H$ be a mapping with a well defined adjoint operator $\mathcal{S}^* : H \times H \mapsto H$. $\mathcal{S}$ is **self-adjoint** if*

$$\mathcal{S}^*(\bar{u}, u)|_{\bar{u}=u} = \mathcal{S}(u), \quad \forall u \in H.$$

Observe that an operator like $\mathcal{T}^*\mathcal{T}(u) := (\mathcal{T}^*\mathcal{T})_1(\bar{u}, u)|_{\bar{u}=u} = \mathcal{T}^*(\mathcal{T}(u), u)$ is always self-adjoint since one may write in terms of the 1st partial adjoint

$$\begin{aligned} \langle \mathcal{T}^*(\mathcal{T}(u), u), \bar{u}\rangle_H &= \langle \mathcal{T}(u), (\mathcal{T}^*)^{*_1}(\bar{u}, \mathcal{T}(u), u)\rangle_H \\ &= \langle u, \underbrace{\mathcal{T}^*((\mathcal{T}^*)^{*_1}(\bar{u}, \mathcal{T}(u), u), u)}_{(\mathcal{T}^*\mathcal{T})^*(\bar{u}, u)}\rangle_H, \end{aligned}$$

or in terms of the 2nd partial adjoint

$$\langle \mathcal{T}^*(\mathcal{T}(u), u), \bar{u}\rangle_H = \langle u, \underbrace{(\mathcal{T}^*)^{*_2}(u, \mathcal{T}(u), \bar{u})}_{(\mathcal{T}^*\mathcal{T})^*(\bar{u}, u)}\rangle_H.$$

In either case, the identities in Theorem 2.3 yield the required property:

$$(\mathcal{T}^*\mathcal{T})^*(\bar{u}, u)|_{\bar{u}=u} = (\mathcal{T}^*\mathcal{T})(u).$$

Example 2.3 Reconsider Example 2.1 where now m=1. It follows that

$$\mathcal{T}^*(y, u)|_{y=u} = uy|_{y=u} = u^2 = \mathcal{T}(u).$$

So $\mathcal{T}$ is self-adjoint. □

Example 2.4 Reconsider Example 2.2 where m=p=1. Even in this SISO case, the corresponding Hammerstein operator is rarely self-adjoint since:

$$\begin{aligned} \mathcal{S}^*(y, u)|_{y=u} &= F(u)S_L^*(u) \\ &\neq \mathcal{S}(u). \end{aligned}$$

□

The final property under consideration is the "C^*-inequality" (iv). Unlike the linear case, only an inequality can relate the two norms in question.

Theorem 2.4 *(C^*-inequality) Let H_1 and H_2 be Hilbert spaces. Assume $\mathcal{T} : H_1 \mapsto H_2$ is a bounded operator with a well-defined adjoint operator. Then the following inequality holds:*

$$\|\mathcal{T}\|^2 \leq \|\mathcal{T}^*\mathcal{T}\|.$$

Proof: For any fixed $u \in H_1$ and employing the Schwarz inequality,

$$\begin{aligned} \|\mathcal{T}(u)\|_{H_2}^2 &= \langle \mathcal{T}(u), \mathcal{T}(u)\rangle_{H_2} \\ &= \langle \mathcal{T}^*(\mathcal{T}(u), u), u\rangle_{H_1} \\ &= \langle \mathcal{T}^*\mathcal{T}(u), u\rangle_{H_1} \\ &\leq \|\mathcal{T}^*\mathcal{T}(u)\|_{H_1}\|u\|_{H_1} \end{aligned}$$

Dividing both sides by $\|u\|_{H_1}^2$ and taking the supremum over all $u \neq 0$ gives the final result. ∎

The section is concluded by considering how the Fréchet derivative interacts with nonlinear Hilbert adjoints. This is important because of its relationship to the eigen-structure of the Hankel operator described in Section 3. Given an operator $\mathcal{U}$ of the form (4), its Fréchet derivative with respect to u_j at $(u_1, u_2, \ldots, u_m)$ is denoted by $D_j\mathcal{U}(u_1, u_2, \ldots, u_m)$. The situation is greatly simplified by the fact that $D_j\mathcal{U}(u_1, u_2, \ldots, u_m)$ is a linear operator defined on H_j.

Theorem 2.5 *Let H_1 and H_2 be two Hilbert spaces and $\mathcal{T} : H_1 \mapsto H_2$ be an operator with a well defined Hilbert adjoint. Assuming both $\mathcal{T}$ and $\mathcal{T}^*$ are Fréchet differentiable, then the following identities hold:*

1. $(D_1\mathcal{T}^*(y, u))^*(u) = \mathcal{T}(u)$
2. $(D_2\mathcal{T}^*(y, u))^*(u) = (D\mathcal{T}(u))^*(y) - \mathcal{T}^*(y, u)$
3. $(D\mathcal{T}^*\mathcal{T}(u))^*(u) = 2\,(D\mathcal{T}(u))^*(\mathcal{T}(u)) - \mathcal{T}^*\mathcal{T}(u)$.

Proof:

1. For any $u \in H_1$ and $\xi, y \in H_2$ observe that
$$\begin{aligned} D_y\langle \mathcal{T}^*(y,u), u\rangle_{H_1}(\xi) &= D_y\langle y, \mathcal{T}(u)\rangle_{H_2}(\xi) \\ \langle D_1\mathcal{T}^*(y,u)(\xi), u\rangle_{H_1} &= \langle \xi, \mathcal{T}(u)\rangle_{H_2} \\ \langle \xi, (D_1\mathcal{T}^*(y,u))^*(u)\rangle_{H_2} &= \langle \xi, \mathcal{T}(u)\rangle_{H_2}. \end{aligned}$$
2. Similarly, for any $u, \xi \in H_1$ and $y \in H_2$
$$D_u\langle \mathcal{T}^*(y,u), u\rangle_{H_1}(\xi) = D_u\langle y, T(u)\rangle_{H_2}(\xi)$$
or
$$\langle D_2\mathcal{T}^*(y,u)(\xi), u\rangle_{H_1} + \langle \mathcal{T}^*(y,u), \xi\rangle_{H_1} = \langle y, D\mathcal{T}(u)(\xi)\rangle_{H_2}$$
and thus,
$$\langle \xi, (D_2\mathcal{T}^*(y,u))^*(u)\rangle_{H_1} = \langle \xi, (D\mathcal{T}(u))^*(y)\rangle_{H_2} - \langle \xi, \mathcal{T}^*(y,u)\rangle_{H_1}.$$
3. First observe that for any $u, \xi \in H_1$
$$\begin{aligned} &\langle u, D\left(\mathcal{T}^*(\mathcal{T}(u), u)\right)(\xi)\rangle_{H_1} = \\ &\langle u, D_1\left(\mathcal{T}^*(\mathcal{T}(u), u)\right)(D\mathcal{T}(u)(\xi)) + \\ &D_2\left(\mathcal{T}^*(\mathcal{T}(u), u)\right)(\xi)\rangle_{H_1}. \end{aligned}$$
Now, apply the previous two identities. ∎

3. Towards Hankel Singular Value Analysis

State space notions have provided useful tools in the case of the nonlinear Hankel operator, e.g., [7, 8]. Consider a smooth time-invariant input-affine nonlinear control system

$$\Sigma : \left\{ \begin{aligned} \dot{x} &= f(x) + g(x)u \\ y &= h(x), \end{aligned} \right. \tag{5}$$

where $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^p$, and $x(t) \in W \subseteq \mathbb{R}^n$ is in local coordinates for a smooth state space manifold. Throughout it is assumed that the system has an isolated equilibrium at 0 and $h(0) = 0$. It is necessary that the system be well defined on the time interval $(-\infty, \infty)$ and that

(A1) Σ is L_2-stable in the sense that $u \in L_2^m(-\infty, 0]$ implies that $\Sigma(u)$ restricted to $[0, \infty)$ is in $L_2^p[0, \infty)$.

The observability and controllability operators for Σ are given by:

$$y = \mathcal{O}_\Sigma(x^0) : \left\{ \begin{aligned} \dot{x} &= f(x) \quad x(0) = x^0 \\ y &= h(x) \end{aligned} \right.$$

$$x^1 = \mathcal{C}_\Sigma(u) : \left\{ \begin{aligned} \dot{x} &= f(x) + g(x)\mathcal{F}_-(u) \quad x(-\infty) = 0 \\ x^1 &= x(0), \end{aligned} \right.$$

where $\mathcal{F}_- : L_2^m(-\infty, \infty) \to L_2^m(-\infty, 0]$ and $\mathcal{F}_+ : L_2^m(-\infty, \infty) \to L_2^m[0, \infty)$ denote the time-flipping operators defined by

$$\mathcal{F}_-(u)(t) := \left\{ \begin{array}{ll} u(-t) & t \in (-\infty, 0] \\ 0 & t \in [0, \infty) \end{array} \right.$$

$$\mathcal{F}_+(u)(t) := \left\{ \begin{array}{ll} 0 & t \in (-\infty, 0] \\ u(-t) & t \in [0, \infty). \end{array} \right.$$

The Hankel operator $\mathcal{H}_\Sigma : L_2^m[0, \infty) \to L_2^p[0, \infty)$ of Σ is given by $\mathcal{H}_\Sigma := \Sigma \circ \mathcal{F}_-$, and the identity $\mathcal{H}_\Sigma = \mathcal{O}_\Sigma \circ \mathcal{C}_\Sigma$ was also proven in [10]. In [8] the use of the adjoint of the variational version of the Hankel operator has been studied and turned out to be useful for an eigen-structure analysis of the Hankel operator. The latter results are summarized next. In order to describe an eigen-structure of the Hankel operator, a state space realization and corresponding pair of energy functions are employed as described below.

Definition 3.1 *The observability function $L_o(x)$ and the controllability function $L_c(x)$ of Σ in (5) are defined by*

$$\begin{aligned} L_o(x^0) &:= \frac{1}{2}\int_0^\infty \|y(t)\|^2 dt, \; x(0) = x^0, \; u(t) \equiv 0 \\ L_c(x^1) &:= \min_{\substack{u \in L_2^m[0,\infty) \\ x(-\infty)=0, x(0)=x^1}} \frac{1}{2}\int_{-\infty}^0 \|u(t)\|^2 dt. \end{aligned}$$

It is assumed throughout that

(A2) there exist well-defined smooth observability and controllability functions L_o and L_c.

Assuming that Σ is Fréchet differentiable, and that $D\Sigma$ is L_2 input-output stable, then the following lemma was proven.

Lemma 3.1 [8] *If there exist $\lambda \in \mathbb{R}$ and a nonzero $x^0 \in \mathbb{R}^n$ such that*

$$\frac{\partial L_o}{\partial x}(x^0) = \lambda\, \frac{\partial L_c}{\partial x}(x^0),$$

then λ is the eigenvalue of the mapping $u \mapsto (D\mathcal{H}_\Sigma(u))^* \circ \mathcal{H}_\Sigma(u)$ with corresponding eigenvector

$$v = \mathcal{C}_\Sigma^\dagger(x^0),$$

where $\mathcal{C}_\Sigma^\dagger : \mathbb{R}^n \to L_2^m[0, \infty)$ denotes the pseudo-inverse of $\mathcal{C}_\Sigma$ defined by

$$\mathcal{C}_\Sigma^\dagger(x^1) := \arg \min_{\mathcal{C}_\Sigma(u)=x^1} \|u\|_2.$$

This lemma relates the gradient of the controllability and observability functions to the eigenvalues of $(D\mathcal{H}_\Sigma(u))^* \circ \mathcal{H}_\Sigma(u)$. The next result gives a more general parameterized eigen-structure of $(D\mathcal{H}_\Sigma(u))^* \circ \mathcal{H}_\Sigma(u)$ in terms of energy level sets in the state space and relates it to $\mathcal{H}_\Sigma^*(\mathcal{H}_\Sigma(u), u)$.

Theorem 3.1 **[8]** *Suppose the energy functions $L_o(x)$ and $L_c(x)$ are sufficiently smooth and that the Jacobian linearization of the system Σ has n distinct Hankel singular values. Then there exists locally $2n$ smooth singular value functions $\rho_i^j(c)$'s, $i \in \{1, 2, \dots, n\}$, $j \in \{+, -\}$ such that $\min\{\rho_i^+(c), \rho_i^-(c)\} > \max\{\rho_{i+1}^+(c), \rho_{i+1}^-(c)\}$, ($\rho_{\max}(c) := \max\{\rho_1^+(c), \rho_1^-(c)\}$, $\rho_{\min}(c) := \min\{\rho_n^+(c), \rho_n^-(c)\}$), and there exists parameterized vectors $x_i^j(c)$'s satisfying*

$$\begin{aligned} L_c(x_i^j(c)) &= \frac{c^2}{2}, \quad L_o(x_i^j(c)) = \frac{c^2(\rho_i^j(c))^2}{2} \\ \frac{\partial L_o}{\partial x}(x_i^j(c)) &= \lambda_i^j(c) \frac{\partial L_c}{\partial x}(x_i^j(c)) \end{aligned} \tag{6}$$

with $\lambda_i^j(c) := (\rho_i^j(c))^2 + (d(\rho_i^j(c))^2/dc)c/2$. Furthermore, when $u_i^j(c) := \mathcal{C}_\Sigma^\dagger(x_i^j(c))$, it follows that

$$\left\langle u_i^j(c), \mathcal{H}_\Sigma^*(\mathcal{H}_\Sigma(u_i^j(c)), u_i^j(c)) \right\rangle_{L_2^m} = (\rho_i^j(c))^2 \left\langle u_i^j(c), u_i^j(c) \right\rangle_{L_2^m},$$

and thus, the Hankel norm of the system is given by

$$\|\Sigma\|_H := \|\mathcal{H}_\Sigma\|_{L_2^m} = \sup_{u \in L_2^m(-\infty, 0]} \frac{\|\mathcal{H}_\Sigma(u)\|_2}{\|u\|_2} = \sup_{c>0} \{\max\{\rho_1^+(c), \rho_1^-(c)\}\}.$$

Spectral theory for nonlinear operators is a diverse subject with substantial roots going back to at least the late 1960's [3]. The proliferation of definitions and approaches (see, for example, [1, 6, 9, 12, 13]) is partly due to the fact that no single definition completely characterizes the original operator as in the linear case. Here we outline an additional approach to defining a nonlinear spectrum motivated by the nature of our application and the notion of the C^1-spectrum introduced in [1].

Definition 3.2 *Let E be a Banach space and $S : E \to E$ be an operator that is continuously Fréchet differentiable on E. The* ***C^1-spectrum*** *of S, $\sigma^1(S)$, is the set of all complex numbers λ such that $S - \lambda I$ is not a diffeomorphism on E.*

For a linear operator S, this definition reduces to the usual definition of a spectrum. The following result is relevant in our study:

Theorem 3.2 [1] *Let S be an operator as described in Definition 3.2, then*

$$\sigma^1(S) = \pi(S) \cup \bigcup_{u \in E} \sigma(DS(u))$$

where $\pi(S)$ denotes the set of all λ such that $S - \lambda I$ is not proper, and $\sigma(\mathcal{A})$ denotes the usual spectrum of a bounded linear operator $\mathcal{A}$.

This theorem reveals that the C^1-spectrum of a nonlinear operator directly involves the Fréchet derivative of the operator. Since we are interested in extending singular value definitions into the nonlinear setting, it is the spectrum of the operator $\mathcal{H}_\Sigma^* \mathcal{H}_\Sigma(u)$ and its derivative that are relevant here. The following corollary of Theorem 3.1 and Theorem 2.5 is directly applicable to the problem.

Corollary 3.1 *In the context of Theorem 3.1, the following relation holds:*

$$\left\langle \left(D\mathcal{H}_\Sigma^* \mathcal{H}_\Sigma(u_i^j(c))\right)(u_i^j(c)), u_i^j(c) \right\rangle = \left(2\lambda_i^j(c) - \rho_i^{j\,2}(c)\right) \left\langle u_i^j(c), u_i^j(c) \right\rangle.$$

Proof: Applying Theorem 2.5, property 3, gives directly

$$\begin{aligned} &\left\langle \left(D\mathcal{H}_\Sigma^* \mathcal{H}_\Sigma(u_i^j(c))\right)(u_i^j(c)), u_i^j(c) \right\rangle = \\ &2\left\langle \left(D\mathcal{H}_\Sigma(u_i^j(c))\right) \left(\mathcal{H}_\Sigma(u_i^j(c))\right), u_i^j(c) \right\rangle - \\ &\left\langle \mathcal{H}_\Sigma^* \mathcal{H}_\Sigma(u_i^j(c)), u_i^j(c) \right\rangle. \end{aligned}$$

Then using Theorem 3.1, the result immediately follows. ∎

Note that the above result yields an eigen-equation within an inner product identity. In the general nonlinear setting, it is not possible to (uniquely) extract the eigen-equation from this identity. Therefore, it is necessary to include the inner product structure directly in a spectrum definition.

Definition 3.3 *Let H be a Hilbert space and consider an operator $S : H \to H$. Then the* ***inner product spectrum*** *is defined as*

$$\sigma_{ip}(S) = \{\lambda : \exists p \neq 0 \text{ with } \langle (S - \lambda I)(p), p \rangle_E = 0\}.$$

It follows immediately when S is continuously Fréchet differentiable on H that $\sigma_{ip}(S) \supset \sigma^1(S)$. Furthermore, in the case of a linear operator $S(p) = Ap$ with $A^T = A$ and $H = \mathbb{R}^n$, it is easily verified that $\sigma_{ip}(S) = \text{Range}(\mathcal{R}_S(p))$, where $\mathcal{R}$ is the Rayleigh quotient of S defined as

$$\mathcal{R}_S(p) = \frac{p^T A p}{p^T p}.$$

It is known in this case that $\sigma_{ip}(A) = [\lambda_{min}(A), \lambda_{max}(A)] \subset \mathbb{R}$, where λ_{min} (λ_{max}) denotes the smallest (largest) eigenvalue of A. The obvious extension of the Rayleigh quotient for nonlinear maps is then

$$\begin{aligned} \mathcal{R}_S &: H \to \mathbb{R} \\ &: p \mapsto \frac{\langle p, S(p) \rangle_H}{\langle p, p \rangle_H}, \end{aligned}$$

and it straightforwardly follows that $\sigma_{ip}(S) = \text{Range}(\mathcal{R}_S)$. This Rayleigh quotient is related to the numerical range $W(S, T)$ as defined in [4] for positively homogeneous operators S and T of degree k on the unit sphere. Indeed, if $T = I$, S is positively homogeneous of degree k and $S_1(0)$ is the unit sphere of H, then $\mathcal{R}_S(S_1(0)) = W(S, I)$.

In the case of a compact linear operator $\mathcal{A} : H \to H$, it is known that

$$\sigma_{ip}(\mathcal{A}^* \mathcal{A}) = (0, \tau_1^2],$$

where τ_1 is the largest singular value of $\mathcal{A}$. If $\text{rank}(\mathcal{A}^* \mathcal{A}) = n < \infty$, then this result can be further refined to

$$\sigma_{ip}(\mathcal{A}^* \mathcal{A}) = [\tau_n^2, \tau_1^2],$$

where τ_n is the smallest singular value of $\mathcal{A}$. In the case of the nonlinear system Σ given by (5), with corresponding Hankel operator $\mathcal{H}_\Sigma$, it follows immediately that $(\rho_i^j(c))^2 \in \sigma_{ip}(\mathcal{H}_\Sigma^* \mathcal{H}_\Sigma)$ for all $i \in \{1, 2, \ldots, n\}$ and $j \in \{+, -\}$. Furthermore,

$$\|\Sigma\|_H^2 = \sup_{c>0} \left\{ \max\{{\rho_1^+}^2(c), {\rho_1^-}^2(c)\} \right\} = \sup \{\sigma_{ip}(\mathcal{H}_\Sigma^* \mathcal{H}_\Sigma)\}$$

and

$$\inf_{c>0} \left\{ \min\{{\rho_n^+}^2(c), {\rho_n^-}^2(c)\} \right\} = \inf \{\sigma_{ip}(\mathcal{H}_\Sigma^* \mathcal{H}_\Sigma)\}.$$

Example 3.5 *Consider the following state space system*

$$\Sigma : \begin{cases} \dot{z}_1 = -z_1 + z_1 z_2^2 + u_1\sqrt{2} \\ \dot{z}_2 = -z_2 - z_2^3 + u_2\sqrt{2 - 2z_1^2 + 2z_2^2} \\ y_1 = 2z_1 \\ y_2 = \sqrt{2} z_2, \end{cases}$$

where $z \in W = \{z | z_1^2 < 1\}$. Then it can be shown that the controllability and observability function are given by

$$\tilde{L}_c(z) = \frac{1}{2} z^T z, \qquad \tilde{L}_o(z) = \frac{1}{2} z^T \begin{pmatrix} 2 & 0 \\ 0 & 1 + z_1^2 \end{pmatrix} z.$$

In order to obtain the $\lambda_i^j(\cdot)$'s in (6), one must compute the solutions of:

$$\begin{cases} 2z_1 + z_1 z_2^2 = \lambda(c) z_1 \\ z_2 + z_1^2 z_2 = \lambda(c) z_2 \\ z_1^2 + z_2^2 = c^2. \end{cases}$$

Note that $z \in W$ implies that $c < 1$, so that $z_1^2 = \frac{1}{2}(1 + c^2)$ and $z_2^2 = \frac{1}{2}(c^2 - 1)$ have no real solution. Thus, there remains two possibilities:

$$\begin{cases} x_1^+(c) = \begin{pmatrix} c \\ 0 \end{pmatrix}, \lambda_1^+(c) = 2 \\ x_1^-(c) = \begin{pmatrix} -c \\ 0 \end{pmatrix}, \lambda_1^-(c) = 2 \end{cases}$$

$$\begin{cases} x_2^+(c) = \begin{pmatrix} 0 \\ c \end{pmatrix}, \lambda_2^+(c) = 1 \\ x_2^-(c) = \begin{pmatrix} 0 \\ -c \end{pmatrix}, \lambda_2^-(c) = 1. \end{cases}$$

In [8] it is shown that

$$c\lambda_i^j(c) = \frac{d}{dc} \left({\rho_i^j}^2(c) \frac{c^2}{2} \right),$$

and thus it follows that

$$\begin{aligned} {\rho_1^\pm}^2(c) = 2, \quad {\rho_2^\pm}^2(c) = 1 \quad &\Rightarrow \quad \sigma_{ip}(\mathcal{H}_\Sigma^* \mathcal{H}_\Sigma) = [1, 2] \\ &\Rightarrow \quad \|\Sigma\|_H = \sqrt{2}. \end{aligned}$$

□

Acknowledgment

This research was supported in part by a travel grant from the Netherlands Organization for Scientific Research (NWO), file number R62-554.

References

[1] J. Appell and M. Dörfner, Some spectral theory for nonlinear operators, *Nonlinear Anal., Theory, Methods & Appl.*, vol. 28, pp. 1955–1976, 1997.

[2] J. Batt, Nonlinear compact mappings and their adjoints, *Math. Ann.*, vol. 189, pp. 5–25, 1970.

[3] F. E. Browder, Nonlinear eigenvalue problems and Galerkin approximations, *Bull. Amer. Math. Soc.*, vol. 74, pp. 651–656, 1968.

[4] V. Burýsková, Some properties of nonlinear adjoint operators, *Rocky Mountain J. of Math.*, vol. 28, no. 1, pp. 41-59, 1998.

[5] P. E. Crouch and A. J. van der Schaft, *Variational and Hamiltonian Control Systems*, Lect. Notes Contr. Inf. Sci. No. 101, Springer Verlag, Berlin, 1987.

[6] G. Fournier and M. Martelli, Eigenvectors for nonlinear maps, *Topol. Methods Nonlin. Anal.*, vol. 2, pp. 203–224, 1993.

[7] K. Fujimoto, J. M. A. Scherpen and W. S. Gray, Hamiltonian realizations of nonlinear adjoint operators, *Proc. IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control*, Princeton, New Jersey, U.S.A., pp. 39–44, 2000.

[8] K. Fujimoto and J. M. A. Scherpen, Eigen structure of nonlinear Hankel operators, in *Nonlinear Control in the Year 2000*, Springer Verlag, A. Isidori, F. Lamnabhi-Lagarrigue and W. Respondek, Eds., 2000.

[9] M. Furi, M. Martelli and A. Vignoli, Contributions to the spectral theory for nonlinear operators in Banach spaces," *Ann. Mat. Pura Appl.*, vol. 118, pp. 229–294, 1978.

[10] W. S. Gray and J. M. A. Scherpen, Hankel operators and Gramians for nonlinear systems, *Proc. 37th IEEE Conf. on Decision and Control*, Tampa, Florida, U.S.A., pp. 1416-1421, 1998.

[11] —, On the nonuniqueness of balanced nonlinear realizations, *Proc. 38th IEEE Conf. on Decision and Control*, Phoenix, Arizona, U.S.A., pp. 4736–4741, 1999.

[12] P. Guillaume, Nonlinear eigenproblems, *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 3, pp. 575–595, 1999.

[13] S. Jingxian and L. Bendong, Eigenvalues and eigenvectors of nonlinear operators and applications, *Nonlin. Anal. Theory Methods Appl.*, vol. 29, no. 11, pp. 1277–1286, 1997.

[14] G. I. Marchuk, V. I. Agoshkov and V. P. Shutyaev, *Adjoint Equations and Perturbation Algorithms in Nonlinear Problems*, CRC Press, Boca Raton, Florida, 1996.

[15] J. M. A. Scherpen, Balancing for nonlinear systems, *Systems & Control Letters*, vol. 21, pp. 143–153, 1993.

[16] J. M. A. Scherpen and W. S. Gray, On singular value functions and Hankel operators for nonlinear systems, *Proc. 1999 IEEE American Control Conf.*, San Diego, California, U.S.A., pp. 2360–2364, 1999.

[17] J. M. A. Scherpen and A. J. van der Schaft, Normalized coprime factorizations and balancing for unstable nonlinear systems, *Int. J. of Contr.*, vol. 60, no. 6, pp. 1193-1222, 1994.